The rough guide to systematic reviews and meta-analyses

G. Biondi-Zoccai[1,2], M. Lotrionte[3], G. Landoni[4], M.G. Modena[1]

[1]*Division of Cardiology, University of Modena and Reggio Emilia, Modena, Italy;*
[2]*Meta-analysis and Evidence-based medicine Training in Cardiology (METCARDIO), Ospedaletti, Italy;*
[3]*Heart Failure and Cardiac Rehabilitation Unit, Catholic University of the Sacred Heart, Rome, Italy;*
[4]*Department of Anesthesia and Intensive Care, Università Vita-Salute San Raffaele, Milan, Italy*

ABSTRACT

The hierarchy of evidence based medicine postulates that systematic reviews of homogenous randomized trials represent one of the uppermost levels of clinical evidence. Indeed, the current overwhelming role of systematic reviews, meta-analyses and meta-regression analyses in evidence based heath care calls for a thorough knowledge of the pros and cons of these study designs, even for the busy clinician. Despite this sore need, few succinct but thorough resources are available to guide users or would-be authors of systematic reviews. This article provides a rough guide to reading and, summarily, designing and conducting systematic reviews and meta-analyses.

Keywords: *meta-analysis, meta-regression, systematic review.*

"I like to think of the meta-analytic process
as similar to being in a helicopter.
On the ground individual trees
are visible with high resolution.
This resolution diminishes as the helicopter
rises, and in its place we begin
to see patterns not visible from the ground"

Ingram Olkin

INTRODUCTION

Systematic reviews and meta-analyses are being used more extensively by researchers and practitioners, thanks to the appeal of a

single piece of literature that is immediately able to summarize diverse data on a specific topic (1, 2). They have been established as the most quoted and read article types, even toppling randomized clinical trials, and thus they are likely to play a progressively even greater role in the future of medicine (3, 4). In addition, they are often published in the most prestigious international peer-reviewed journals, reaching thousands of physicians and researchers worldwide.

As with any other analytical and research tool with a long-standing history (*Table 1*), systematic reviews and meta-analyses, despite their major strengths, are well known for several potential major weaknesses.

The aim of this review is to provide a concise but sound framework for the critical reading of systematic reviews and meta-analyses and, summarily, their design and

Corresponding author:
Giuseppe Biondi-Zoccai, MD
Division of Cardiology,
University of Modena and Reggio Emilia,
Via Del Pozzo, 71 - 41124 Modena, Italy
email: gbiondizoccai@gmail.com

Table 1 - *Key milestones in systematic review and meta-analysis development.*

Year	Individuals	Milestone
1904	Karl Pearson (UK)	correlation between inoculation of vaccine for typhoid fever and mortality across apparently conflicting studies
1931	Leonard Tippet (UK)	comparison of differences between and within farming techniques on agricultural yield adjusting for sample size across several studies
1937	William Cochran (UK)	combination of effect sizes across different studies of medical treatments
1970s	Robert Rosenthal, Gene Glass (USA); Archie Cochrane (UK)	combination of effect sizes across different studies of, respectively, educational/psychological and clinical treatments
1980s	The global scientific community	exponential development/use of meta-analytic methods; birth of The Cochrane Collaboration

conduct, stemming from our extensive experience with this type of research method (*Figure 1*).

Definitions

A systematic review is a viewpoint focusing on a specific clinical problem, being it therapeutic, diagnostic or prognostic (*Table 2*) (1, 5). The term systematic means that all the steps underlying the reviewing process are explicitly and clearly defined, and may be reproduced independently by other researchers. Thus, a formal set of methods is applied to study search (i.e. to the extensive search of primary/original studies), study selection, study appraisal, data abstraction and, when appropriate, data pooling according to statistical methods. Indeed, the term meta-analysis refers to a statistical method used to combine results from several different primary studies in order to provide more precise and valid results. Thus, not all systematic reviews include a meta-analysis, as not all topics are
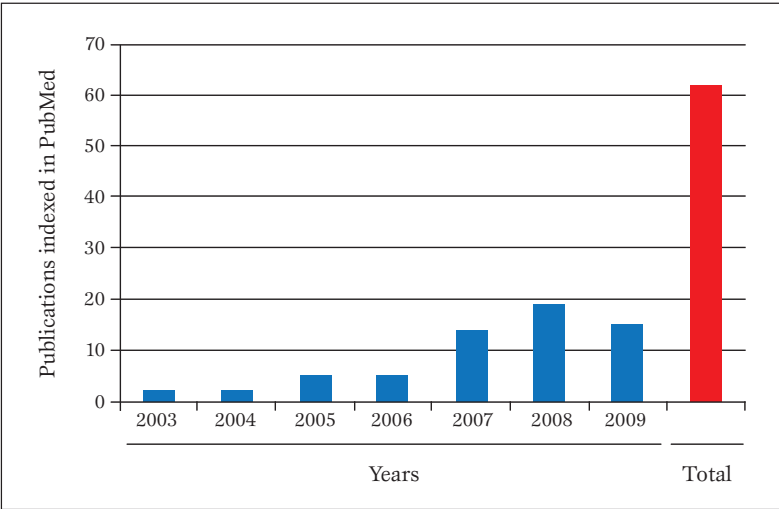

Figure 1 - *Publications in PubMed authored in the last few years by our research group concerning meta-analytic topics. PubMed was searched on 30 March 2010 with the following strategy: "(biondi-zoccai OR Zoccai) AND (meta-analys* OR metaanalys* OR metaregress* OR "meta-regression")".*

Table 2 - *Minimal glossary pertinent to systematic reviews and meta-analyses.*

Term	Characteristics
Review	A viewpoint on a given subject quoting different primary authors or studies
Overview	As above
Qualitative review	A review which avoids a systematic approach
Systematic review	A review which deliberately exploits and report a systematic approach to study search, selection, abstraction, appraisal and pooling
Quantitative review	A review which deliberately exploits and report quantitative methods to evaluate or synthesize data
Meta-analysis	A study (not necessarily a review) using specific statistical methods for pooling data from separate datasets
Meta-regression	A study (not necessarily a review) using specific statistical methods for exploring interactions between dependent and independent variables (moderators) from a meta-analysis dataset
Individual patient data meta-analysis	A study (not necessarily a review) using specific statistical methods for pooling data from separate datasets exploiting individual patient data
Overview of reviews	An overview of reviews which deliberately a systematic approach to review search, selection, abstraction, appraisal and pooling

suitable for sound and robust data pooling. At the same time, meta-analysis can be conducted outside the realm of a systematic review (e.g. in the absence of extensive and thorough literature searches), but in such cases results of the meta-analytic efforts should be best viewed as hypothesis-generating only. This depends mainly on the fact that meta-analysis outside the framework of a systematic review has a major risk of publication bias.

Strenghts

Systematic reviews (especially when including meta-analytic pooling of quantitative data) have several unique strengths (1, 5). Specifically, they exploit systematic literature searches enabling the retrieval of the whole body of evidence pertaining to a specific clinical question.

Their standardized methods for search, evaluation and selection of primary studies enable reproducibility and an objective stance. Individual primary studies undergo a thorough evaluation for internal validity, together with the identification of the risk for bias All too often, systematic reviews hold their greater strength precisely in their ability to pinpoint weaknesses and fallacies in apparently sound primary studies (6).

Quantitative synthesis by means of meta-analysis also substantially increases statistical power, and yields narrower confidence intervals for statistical inference. The assessment of the effect of an intervention (exposure or diagnostic test) across different settings and times provides estimates and inferences with much greater external validity. The larger sample sizes often achieved by systematic reviews may even offer ample room for testing post-hoc hypotheses or exploring the effects in selected subgroups (7). Clinical and statistical variability (i.e. heterogeneity and inconsistency) may be exploited by advanced statistical methods such as meta-regression, possibly offering the opportunity to test novel and hitherto unprecedented hypotheses (8). Fi-

nally, meta-regression methods can be used to perform adjusted indirect comparisons or network meta-analyses (9).

Limitations

Drawbacks of systematic reviews and meta-analyses are also substantial, and should never be dismissed (1). Since the first critique of being "an exercise in mega-silliness" and inappropriately "mixing apples and oranges"(10), there has been ongoing debate on the most correct approach to choose when meta-analytic pooling should be pursued (e.g. in case of statistical homogeneity and consistency) and when, conversely, the reviewer should refrain from meta-analysis (e.g. in case of severe statistical heterogeneity [as testified by p values < 0.10 at χ^2 test] or significant statistical inconsistency [as testified by I^2 values > 50 %]) (11).

Whereas Canadian authors suggest that systematic reviews and meta-analyses from homogenous randomized controlled trials represent the apex of the evidence-based medicine pyramid (discounting for the role of n of 1 randomized trials) (12), others maintain that very large and sim-

ple randomized clinical trials offer several premium features, and should always be preferred, when available, to systematic reviews (13).

It is also all too common to retrieve only a few studies which focus on a given clinical topic, or otherwise studies may be found, but of such low quality, that including or even discussing them in the setting of a systematic review may appear misleading. Indeed, in such cases the meta-analysis itself can be considered misleading. Nonetheless, key insights may be gained in these cases by exploring sources of heterogeneity, stratified analyses, and meta-regressions. This drawback is strictly associated with the major threat to meta-analysis validity called the small study effect (also, albeit inappropriately, called small study bias or publication bias) (1). Indeed, it is common to recognize, especially in large datasets, that small primary studies are more likely to be reported, published and quoted if their results are significant. Conversely, small non-significant studies often fail to reach publication or dissemination, and may thus be very easily missed, even after thorough
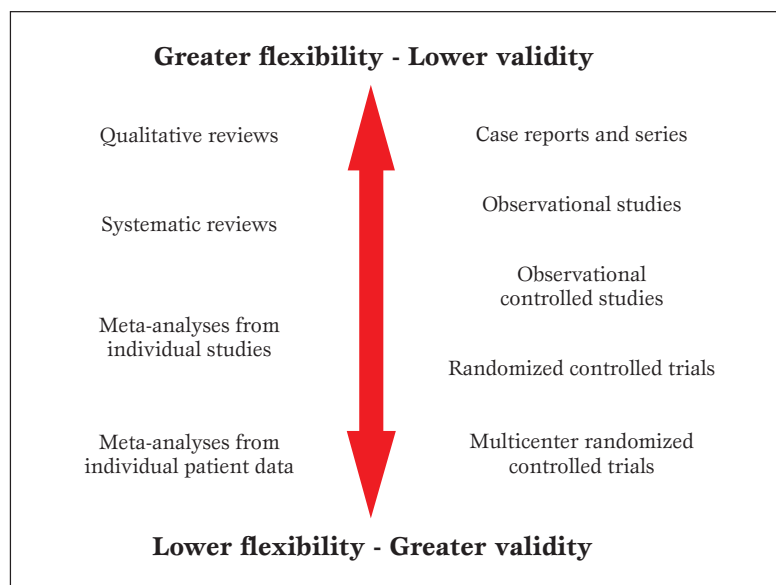
Figure 2 - *Parallel hierarchy of scientific studies in clinical research. Modified from Biondi-Zoccai et al.(2).*

literature searches. Combining results from these "biased" small studies with those of larger studies (which are usually published even if negative or non-significant) may inappropriately deviate summary effect estimates away from the true value. Unfortunately, despite the availability of several graphical and analytical tests (14) small study effects (which actually encompass publication bias) are potentially always present in a systematic review and should never be disregarded.

In addition, in the ongoing worldwide research effort, it is all too common for reviewers to focus only on English language studies, and thus unduly restricting their search and excluding potentially important works (e.g. from China or Japan). Another common critique is that systematic reviews and meta-analyses are not original research (*Figure 2*). The reader is left free to form independently his informed opinion on this specific issue. Nonetheless, the main meter to judge a systematic review should be its novelty and usefulness for the very same reader, not whether it appears as original or secondary research (2).

Finally, a burning issue is whether results from large systematic reviews and meta-analyses can ever be applied to the single individual under our care. This question cannot be answered once and for all, and judgment should always be employed when considering the application of meta-analytic results to a specific patient. Unless proven otherwise by a significant test of interaction, all patients should be considered likely to similarly benefit from a specific treatment or diagnostic strategy (12).

Appraising primary studies, systematic reviews and meta-analyses

Unfortunately publication of a systematic review in a peer-reviewed journal is not definitive evidence of its internal validity and usefulness for the clinical practitioner or researcher (15). Peer-review is not very accomplished in judging or improving the quality of scientific reviews, and many examples of bad or unsuccessful peer-reviewing efforts can be easily found. However, just as "democracy is the worst form of government except all those other forms that have been tried" (Sir Winston Churchill), peer-review is the "worst" method used to evaluate scientific research except all other methods that have been tried so far. This applies to all clinical research products in general and so also applies to systematic reviews and meta-analyses. Thus, provided that meta-analyses are accurately and thoroughly reported, the burden of quality appraisal lies largely, as usual, in the eye of the beholder (i.e. the reader).

Assessment of primary research studies as well as systematic reviews and meta-analyses should be based on their internal validity and then, provided it is reasonably adequate, on their results and external validity (12). Whereas interpretation of results and external validity of any research endeavor depends on the specific context of application, and is thus best left open to the individual judgment of the reader or decision-maker, internal validity can be evaluated in a rather structured and validated way. Recent guidance on the appraisal of the risk of bias in primary research studies within the context of a systematic review has been provided by The Cochrane Collaboration, and includes a separate assessment of the risk of selection, performance, attrition and adjudication bias (*Table 3*) (16). Other valid and complementary approaches, targeted for specific study designs, have been proposed by advocates of evidence-based medicine methods, and include the Jadad score, the Delphi list, and the Megens-Harris list (12). Nonetheless, even external validity can be formally evaluated by focusing on the population included, the control group, and result interpretation. Finally,

Table 3 - *A modified version of The Cochrane Collaboration risk of bias assessment tool for the appraisal of primary studies.(16) **

Question	Answers	Meaning
Adequate sequence generation?	Yes, no, or uncertain	Was the allocation sequence generated appropriately (eg computer or table of random numbers)?
Allocation concealment used?	Yes, no, or uncertain	Were physicians unaware of allocation code up to actual patient enrolment?
Blinding?	Yes, no, or uncertain	Were patients, caregivers, outcome assessors, ancillary personnel, and/or statisticians unaware of actual treatment?
Concurrent therapies similar?	Yes, no, or uncertain	Were concurrent medical and non-medical treatments similar in the groups under comparison?
Incomplete outcome data addressed?	Yes, no, or uncertain	Were all data analyzed, minimizing the impact of losses to follow-up?
Uniform and explicit outcome definitions?	Yes, no, or uncertain	Were definitions clearly spelled out and employed consistently to adjudicate events or outcomes?
Free of selective outcome reporting?	Yes, no, or uncertain	Were all relevant outcomes thoroughly reported?
Free of other bias?	Yes, no, or uncertain	Was the risk of any other bias low?
Overall risk of bias?	High, moderate, or low	What is the comprehensive assessment of the risk of bias of the study?

*The following items are not present in the original version supported by The Cochrane Collaboration: concurrent therapies similar; uniform and explicit outcome definitions; overall risk of bias

Table 4 - *Oxman and Guyatt index for the appraisal of reviews.(19) **

Question	Details
1	Were the search methods used to find evidence stated?
2	Was the search for evidence reasonably comprehensive?
3	Were the criteria for deciding which studies to include in the overview reported?
4	Was bias in the selection of studies avoided?
5	Were the criteria used for assessing the validity of the included studies reported?
6	Was the validity of all studies referred to in the text assessed using appropriate criteria
7	Were the methods used to combine the findings of the relevant studies reported
8	Were the findings of the relevant studies combined appropriately relative to the primary question the overview addresses?
9	Were the conclusions made by the author(s) supported by the data and/or analysis reported in the overview?
10	This question summarizes the previous ones and, specifically, asks to rate the scientific quality of the review from 1 (being extensively flawed) to 3 (carrying major flaws) to 5 (carrying minor flaws) to 7 (minimally flawed)

*The Oxman and Guyatt index evaluates the internal validity of a review on 9 separate questions for which 3 distinct answers are eligible ("yes", "partially/can't tell", "no"). The developers of the index specify that if the "partially/can't tell" answer is used one or more times in questions 2, 4, 6, or 8, a review is likely to have minor flaws at best and is difficult to rule out major flaws (ie a score≤4). If the "no" option is used on question 2, 4, 6 or 8, the review is likely to have major flaws (ie a score≤3)

established statistical criteria are available to determine whether a given intervention is effective and similar explicit criteria can inform on the presence of clinical significance.

The quality of a systematic review and meta-analysis depends on several factors, in particular the quality of the primary pooled studies.

Nonetheless, reporting quality (e.g. compliance with current guidelines on drafting and reporting of a meta-analysis by the Preferred Reporting Items for Systematic reviews and Meta-Analyses [PRISMA] or Meta-analysis Of Observational Studies in Epidemiology [MOOSE] statements) should be clearly distinguished by internal validity (17, 18). This can be low even in well reported reviews, whereas it is generally difficult to judge as highly valid a poorly reported systematic review and meta-analysis. The assessment of the internal validity of a review is quite complex and based on several assumptions, including study search and appraisal, methods for data pooling, and approaches to interpretation of study findings.

However, useful guidance was provided by Oxman and Guyatt with their well validated instrument (*Table 4*) (19).

More recently, other investigators have suggested other tools for the evaluation of systematic reviews, such as the A Measurement Tool to Assess Systematic Reviews (AMSTAR), and the Veritas plot, which await further validation (*Table 5, Figure 3*) (20-22).

For those busy critical care physicians wishing for a quicker approach to appraise systematic reviews, a simple two-step approach can be proposed. This is a simplification of the evidence-based medicine approach for the evaluation of sources of clinical evidence, but is nonetheless quite helpful (12). Evidence-based medicine is "the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients" (12). It must also be stressed that "the practice of evidence-based medicine requires integration of individual clinical expertise and patient references with the best available external clinical evidence from systematic search" (12). Systematic reviews

Table 5 - *The AMSTAR tool for the appraisal of systematic reviews.(20-21) **

Question	Details
1	Was an 'a priori' design provided?
2	Was there duplicate study selection and data extraction?
3	Was a comprehensive literature search performed?
4	Was the status of publication (i.e. grey literature) used as an inclusion criterion?
5	Was a list of studies (included and excluded) provided?
6	Were the characteristics of the included studies provided?
7	Was the scientific quality of the included studies assessed and documented?
8	Was the scientific quality of the included studies used appropriately in formulating conclusions?
9	Were the methods used to combine the findings of studies appropriate?
10	Was the likelihood of publication bias assessed?
11	Was the conflict of interest stated?

*The AMSTAR (a measurement tool to assess the methodological quality of systematic reviews), evaluates the quality of a review on 11 separate questions for which 4 distinct answers are eligible ("yes", "no", "can't answer", "not applicable")

and meta-analyses, if well conducted and reported, help us in reducing our efforts in looking for, evaluating, and summarizing the evidence.

But the burden of deciding what to do with the evidence obtained for the care of our individual patient remains ours.

Thus, the first step in appraising a system-

atic review and meta-analysis is to try and find an answer to the question: can I trust it? In other words, is this review internally valid, does it provide a precise and largely unbiased answer to its scientific question? Providing a definitive assessment of the internal validity of a systematic review is not a simple task, but largely depends on
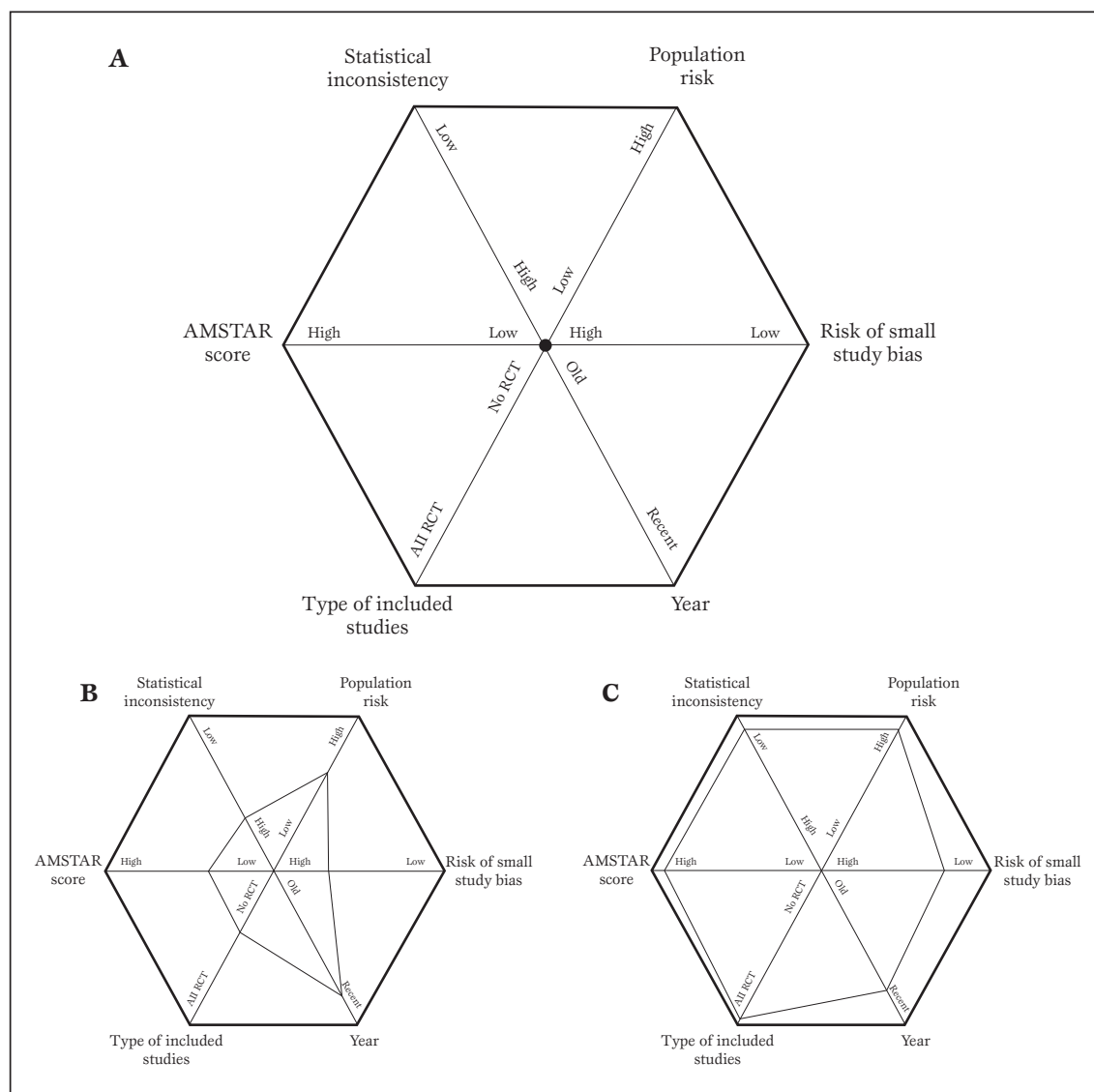

Figure 3 - *Typical diagram used to generate a Veritas plot (panel A) (22).*
Using this tool, a low quality meta-analysis will be represented by a hexagon with a smaller area (panel B), whereas a high quality meta-analysis will be shown as a hexagon with a larger area (panel C).

the methods employed and reported regarding study search, selection, abstraction, appraisal and, if appropriate, the study pooling. Even if we can conclude that a given meta-analysis is internally valid, we still have to face the second step in its evaluation. This focuses on the external validity of the study. In other words, can I apply the review results to the case I am facing or will shortly face? More basically it answers the question: so what? Decisions on external validity are highly subjective and may change depending on the clinical, historical, logistical, cultural or ethical context of the evaluator. Nonetheless, systematic reviews and meta-analyses can improve our appraisal of the external validity of any given clinical intervention, by suggesting an overall clinical efficacy (or lack of it).

It is clear that the assessment of the internal validity, and even more importantly the external validity, of any research endeavor, is highly subjective, and thus we leave ample room for the reader to enjoy and appraise them on his or her own.

The only issue that is worth being further stressed is that only collective and constructive, but critical post-publication evaluation of scientific studies can put and maintain them into the appropriate context for their correct and practical exploitation

by the clinical researcher and the clinical practitioner.

Systematic reviews and meta-analyses: do it yourself

Even those not strictly committed to conduct a systematic review may obtain further insights into this clinical research method by understanding the key steps involved in the design, conduct and interpretation of a systematic review (5).

Briefly, a systematic review should always stem from a specific clinical question. Even if the experienced reviewer can probably informally guess the answer to this question the goal of the systematic review will be to confirm or disprove such hypothesis in a formal and structured way. With this goal in mind, the review should be designed as prospectively and in as much in detail as possible, to avoid conscious or unconscious manipulations of methods or data (*Figure 4*).

The next steps are very important, and define the boundaries of the reviewing effort. Specifically, the reviewer should spell out the population of interest, the intervention or exposure to be appraised, the comparison(s) or comparator(s), and the outcome(s). The acronym PICO is often used to remember this approach. As an
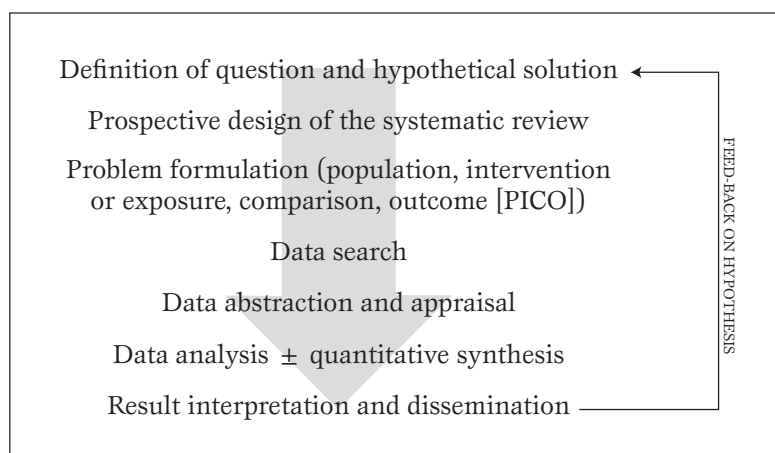

Figure 4 - *Typical algorithm for the design and conduct of a systematic review. Modified from Biondi-Zoccai et al.(5).*

Definition of question and hypothetical solution

Prospective design of the systematic review

Problem formulation (population, intervention or exposure, comparison, outcome [PICO])

Data search

Data abstraction and appraisal

Data analysis ± quantitative synthesis

Result interpretation and dissemination

FEED-BACK ON HYPOTHESIS

example, we could be interested in conducting a systematic review focusing on a population (P) of diabetics with coronary artery disease undergoing coronary artery bypass grafting, with the intervention (I) of interest being the administration of bivaridudin as anticoagulant, the comparator (C) being unfractioned heparin, and the outcomes (O) defined as in-hospital rates of death, myocardial infarction, stroke, or major bleeding (including bleeding needing repeat surgery).

After such preliminary steps, the actual review begins with a thorough and extensive search, encompassing several databases (not only MEDLINE/PubMed) with the help of library personnel experienced in literature searches, preferably also including conference abstracts and bibliographies of pertinent articles and reviews. When a list of potentially pertinent citations has been retrieved, these should be assessed and included/excluded based on criteria

stemming directly from the PICO approach used to define the clinical question. Study appraisal also includes a formal evaluation of study validity and risk of bias of primary studies, whereas data abstraction, generally performed by at least two independent reviewers with divergences resolved after consensus, provides the quantitative data which will eventually be pooled with meta-analysis (16).

Indeed, provided that studies are relatively homogeneous and consistent, meta-analytic methods are employed to combine effect estimates from single studies into a unique summary effect estimate, with corresponding p values and confidence intervals for the effect (*Figure 5*). In many cases results may lead reviewers to go back to the original research question and revise their working hypothesis. The last step relies on the interpretation and dissemination (possibly through publication in a peer-reviewed journal) of the results.
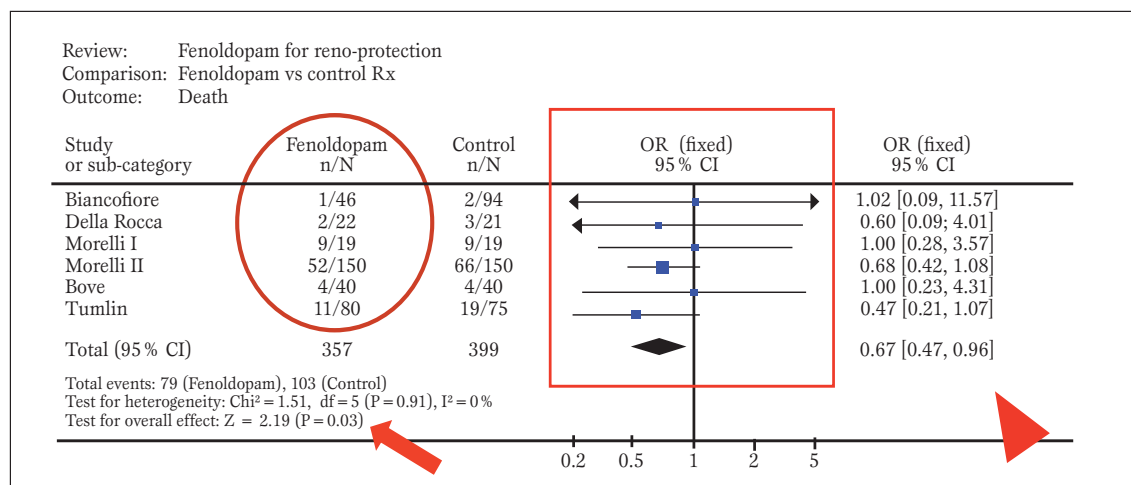

Figure 5 - *Typical forest plot generated by RevMan from a systematic review with meta-analytic pooling of dichotomous outcomes (df = degrees of freedom; E = expected cases; O = observed cases; OR = odds ratio). The solid oval highlights event counts in one of the groups under comparison, the solid box shows graphically individual and pooled point effect estimates with 95% confidence intervals, the arrowhead indicates the exact pooled point effect estimate with 95% confidence intervals (CI), the arrow shows the p value for effect, and the dashed oval highlights p value for statistical heterogeneity and measure of statistical inconsistency (I2). Modified from Landoni et al.(30).*

More advanced analytical issues

Unless extensively powered low event rates may often be found in primary research studies (e.g. with > 1000 patients enrolled or with selective recruitment of very high-risk subjects). This may lead to null counts in one or more of the groups undergoing comparison in a controlled trial, generating severe computational hurdles. Indeed, most statistical methods used for meta-analytic pooling require that at least one event has occurred in each study group.

When this is not the case in one or more of the groups under comparison, bias may be introduced with the common practice of adding 0.25 or 0.50 to each group without events (23). On top of this, when no event has occurred in any group, comparisons are more challenging and data from such an underpowered studies cannot be pooled with standard meta-analytic methods, as variance of the effect estimate approaches infinity. Nonetheless, other approaches (e.g. risk difference, continuity correction or Peto method) can still be used in case of total zero event trials.

Even when all groups undergoing comparison in a specific study have one or more events, the risk of biased estimates and alpha error (i.e. the risk of erroneously dismissing a null hypothesis despite it being true) may be present (1).

Indeed, minor differences in populations with few and rare events may provide nominally significant results (e.g. $p = 0.048$) which however appear quite unstable. In such cases, we recommend reliance on the combined use of p values and 95 % confi-
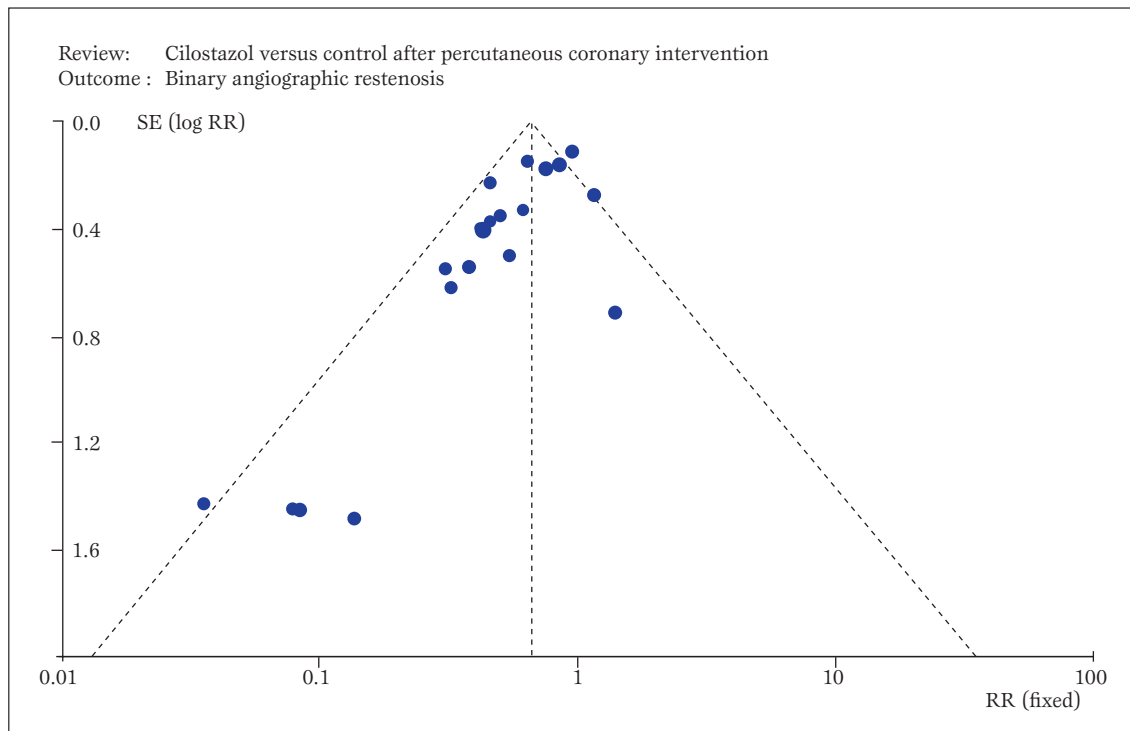

Figure 6 - *Typical funnel plot generated by RevMan showing small study bias, ie the asymmetric distribution of effect sizes in function of study precision, with selective publication of only positive small sample studies (RR = relative risk; SE = standard error). Modified from Biondi-Zoccai et al.(24).*

dence intervals, or even making use of 99 % confidence intervals. In other cases, a useful rule of thumb is to trust only meta-analyses reporting on at least 100 pooled events per group under comparison.

The risk of erroneously accepting a null hypothesis despite it being false (i.e. the beta error) is also common in systematic reviews and meta-analyses, especially when they include few studies with low event counts. This lack of statistical power (defined as 1-beta) is even more common with meta-regression analyses, which are usually underpowered because of few included studies and regression to the mean (7).

Surrogates may provide an important contribution to clinical research design, by increasing statistical power and offering insights in more than one clinical dimension. However, surrogate end-points (e.g. > 25 % increase in serum creatinine from baseline values to identify subclinical renal injury) may be less clinically relevant than hard clinical end-points (death or permanent need for hemodialysis) (12).

Usually, only surrogates which have a direct impact on patient well being and are independently associated with hard clinical end-points should be accepted for the design of clinical research studies. In any case, a study reaching significance based on surrogate end-points alone, but missing significance on analysis of hard end-points should be considered as hypothesis-generating or, at best, underpowered.

Small study bias always potentially threatens the results of a systematic review, as this type of confounding applies to all clinical topics and research study designs (*Figure 6*) (24).

Although this bias may be less significant in more recent and well financed drug or device studies (e.g. fenoldopam), in older or less well funded studies publication bias may profoundly undermine the results of a systematic review.

This has been all too evident in studies examining the role of acetylcysteine for the prevention of contrast-associated nephropathy (25), but is also obvious in other commonly prescribed agents. Another major threat to the validity of a systematic review, as to any other research endeavor, lies in conflicts of interest and study funding. It is well known that reviewers with underlying financial conflicts of interest are more likely to conclude in favor of the intervention benefiting the source of financial gains (26).

Whether these facts should lead to a more critical reading of their work or a comprehensive re-evaluation of their whole research project is best left at the readers' discretion, but this should also take into account the overall internal validity (e.g. blinding of patients, physicians, adjudicators, and analysts) of the work.

CONCLUSIONS

Systematic reviews and meta-analyses offer powerful methods to evaluate the clinical effects of health interventions, especially when directly applied to real world clinical practice (such as in the Best Evidence Topic [BET] approach) (27).

More collaborative efforts are however required to design, conduct and disseminate individual patient data meta-analyses in an unbiased and rigorous manner (28, 29).

REFERENCES

1. Egger M, Smith GD, Altman DG. Systematic reviews in health care: meta-analysis in context, 2nd ed. 2001: BMJ Publishing Group, London.
2. Biondi-Zoccai GG, Agostoni P, Abbate A. Parallel hierarchy of scientific studies in cardiovascular medicine. Ital Heart J 2003; 4: 819-20.
3. Patsopoulos NA, Analatos AA, Ioannidis JP. Relative citation impact of various study designs in the health sciences. JAMA 2005; 293: 2362-6.
4. Glasziou P, Djulbegovic B, Burls A. Are systematic reviews more cost-effective than randomised trials? Lancet 2006; 367: 2057-8.

5. Biondi-Zoccai GG, Testa L, Agostoni P. A practical algorithm for systematic reviews in cardiovascular medicine. Ital Heart J 2004; 5: 486-7.

6. Lau J, Ioannidis JP, Schmid CH. Summing up evidence: one answer is not always enough. Lancet 1998; 351: 123-7.

7. Thompson SG, Higgins JP. How should meta-regression analyses undertaken and interpreted? Stat Med 2002; 21: 1559-73.

8. Biondi-Zoccai GG, Abbate A, Agostoni P, et al. Long-term benefits of an early invasive management in acute coronary syndromes depend on intracoronary stenting and aggressive antiplatelet treatment: a metaregression. Am Heart J 2005; 149: 504-11.

9. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. J Clin Epidemiol 1997; 50: 683-91.

10. Glass G. Primary, secondary and meta-analysis of research. Educ Res 1976; 5: 3-8.

11. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003; 327: 557-60.

12. Guyatt G, Rennie D, Meade M, Cook D. Users' guides to the medical literature. A manual for evidence-based clinical practice. 2002: AMA Press, Chicago.

13. Cappelleri JC, Ioannidis JP, Schmid CH, et al. Large trials vs meta-analysis of smaller trials: how do their results compare? JAMA 1996; 276: 1332-8.

14. Peters JL, Sutton AJ, Jones DR, et al. Comparison of two methods to detect publication bias in meta-analysis. JAMA 2006; 295: 676-80.

15. Opthof T, Coronel R, Janse MJ. The significance of the peer review process against the background of bias: priority ratings of reviewers and editors and the prediction of citation, the role of geographical bias. Cardiovasc Res 2002; 56: 339-46.

16. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions. 2008: The Cochrane Collaboration, Oxford.

17. Liberati A, Altman DG, Tetzlaff J,et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 2009; 339: 2700.

18. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA 2000; 283: 2008-12.

19. Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. J Clin Epidemiol 1991; 44: 1271-8.

20. Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol 2007; 7: 10.

21. Shea BJ, Bouter LM, Peterson J, et al. External validation of a measurement tool to assess systematic reviews (AMSTAR). PLoS One 2007; 2: 1350.

22. Panesar SS, Rao C, Vecht JA, et al. Development of the Veritas plot and its application in cardiac surgery: an evidence-synthesis graphic tool for the clinician to assess multiple meta-analyses reporting on a common outcome. Can J Surg 2009; 52: 137-45.

23. Golder S, Loke Y, McIntosh H. Room for improvement? A survey of the methods used in systematic reviews of adverse effects. BMC Med Res Methodol 2006; 6: 3.

24. Biondi-Zoccai GG, Lotrionte M, Anselmino M, et al. Systematic review and meta-analysis of randomized clinical trials appraising the impact of cilostazol after percutaneous coronary intervention. Am Heart J 2008; 155: 1081-9.

25. Biondi-Zoccai GG, Lotrionte M, Abbate A, et al. Compliance with QUOROM and quality of reporting of overlapping meta-analyses on the role of acetylcysteine in the prevention of contrast associated nephropathy: case study. BMJ 2006; 332: 202-09.

26. Barnes DE, Bero LA. Why review articles on the health effects of passive smoking reach different conclusions. JAMA 1998; 297: 1566-70.

27. Dunning J, Prendergast B, Mackway-Jones K. Towards evidence-based medicine in cardiothoracic surgery: best BETS. Interact Cardiovasc Thorac Surg 2003; 2: 405-9.

28. De Luca G, Gibson CM, Bellandi F, et al. Early glycoprotein IIb-IIIa inhibitors in primary angioplasty (EGYPT) cooperation: an individual patient data meta-analysis. Heart 2008; 94: 1548-58.

29. Burzotta F, De Vita M, Gu YL, et al. Clinical impact of thrombectomy in acute ST-elevation myocardial infarction: an individual patient-data pooled analysis of 11 trials. Eur Heart J 2009; 30: 2193-203.

30. Landoni G, Biondi-Zoccai GG, Tumlin JA, et al. Beneficial impact of fenoldopam in critically ill patients with or at risk for acute renal failure: a meta-analysis of randomized clinical trials. Am J Kidney Dis 2007; 49: 56-68.